

From bright binaries to bumpy backgrounds: Mapping realistic gravitational wave skies with pulsar-timing arrays

Stephen R. Taylor^{1,2,*}, Rutger van Haasteren^{2,3} and Alberto Sesana⁴

¹*Department of Physics & Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, Tennessee 37235, USA*

²*TAPIR Group, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA*

³*Microsoft Corporation, One Microsoft Way, Redmond, Washington 98052, USA*

⁴*Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*



(Received 9 June 2020; accepted 25 August 2020; published 15 October 2020)

Within the next several years, pulsar-timing array programs will likely usher in the next era of gravitational-wave astronomy through the detection of a stochastic background of nanohertz-frequency gravitational waves, originating from a cosmological population of inspiraling supermassive binary black holes. While the source positions will likely be isotropic to a good approximation, the gravitational-wave angular power distribution will be anisotropic, with the most massive and/or nearby binaries producing signals that may resound above the background. We study such a realistic angular power distribution, developing fast and accurate sky-mapping strategies to localize pixels and extended regions of excess power while simultaneously modeling the background signal from the less massive and more distant ensemble. We find that power anisotropy will be challenging to discriminate from isotropy for realistic gravitational-wave skies, requiring $\text{SNR} > 10$ in order to favor anisotropy with 10:1 posterior odds in our case study. Amongst our techniques, modeling the population signal with multiple point sources in addition to an isotropic background provides the most physically motivated and easily interpreted maps, while spherical-harmonic modeling of the square-root power distribution, $P(\hat{\Omega})^{1/2}$, performs best in discriminating from overall isotropy. Our techniques are modular and easily incorporated into existing pulsar-timing array analysis pipelines.

DOI: [10.1103/PhysRevD.102.084039](https://doi.org/10.1103/PhysRevD.102.084039)

I. INTRODUCTION

High-precision timing of millisecond pulsars (MSPs) has been, and continues to be, a valuable tool for probing a wide range of physics, from studies of nuclear matter to tests of modified gravity theories. Successes include the first indirect confirmation of gravitational-wave (GW) emission [1,2], and very accurate tests of general relativity [3,4]. Pulsar-timing array (PTA) projects aim to *directly* detect low-frequency GWs in the range 10^{-9} – 10^{-7} Hz from extra-Galactic sources by using a set of Galactic MSPs as nearly perfect Einstein clocks [5]. This endeavor is possible due to the exceptional regularity of pulses and the remarkable stability of pulse profiles. Many processes must be modeled in a pulsar's timing ephemeris, but if done so accurately one can account for every rotation of the pulsar across observation epochs. The presence of GWs affects the propagation of pulses from the pulsar to the Earth, changing the proper length of the photon path, and

creating detectable deviations away from the expected pulse times of arrival (TOAs) [6–8].

Over the past two decades, several regional collaborations have been collecting data and hunting for a stochastic GW background, periodic GW signals, GW memory bursts, and more; these include the European Pulsar Timing Array (EPTA, [9]) [10–12], the North American Nanohertz Observatory for Gravitational Waves (NANOGrav, [13]) [14–20], and the Parkes Pulsar Timing Array (PPTA, [21,22]) [23–27]. These three were founding members of the International Pulsar Timing Array (IPTA, [28,29]), aimed at strengthening pulsar sky coverage and observation baselines. There are also burgeoning PTA GW search efforts in India (InPTA, [30]), China (CPTA, [31]), as well as in telescope-centered timing groups like MeerTime [32] and CHIME/Pulsar [33]. These ongoing collaborative searches have yielded precision timing data on more than 80 MSPs, some with baselines longer than 15 years.

Binary systems of supermassive black holes (SMBHs) with masses $\sim 10^8$ – $10^{10} M_{\odot}$ are expected to be the dominant source of GWs for PTAs [34–37]. Such titanic black holes lie at the heart of massive galaxies, forming binary systems as

*stephen.r.taylor@vanderbilt.edu

their host galaxies merge and grow over the history of the Universe [38–41]. After the merger of two galaxies, the resident SMBHs will each experience dynamical friction in the merger remnant, carrying them to $\sim 1\text{--}10$ parsec separations to form a bound supermassive binary black hole (SMBBH) system. At this stage, discrete scattering events with central bulge stars will evolve the binary toward even closer separations, possibly giving way to viscous interactions with a circumbinary disk until the binary becomes dominated by GW emission at separations of $\lesssim 10^{-3}$ parsecs (see e.g., [42], and references therein). The system will then be emitting GWs at frequencies that are accessible to PTAs. While PTAs are in principle sensitive to the entire cosmological population of SMBBHs, discriminating each and every one is beyond the resolution limit [43–45]. Hence the first expected signal will be that of the incoherent superposition of all binaries producing a stochastic GW background (GWB) [46–48], above which a few particularly nearby or massive binaries may resound as individually detectable signals [24,43,49–55].

Beyond a few hundred Mpc, the Universe can be approximated as statistically homogeneous and isotropic. Hence the (an)isotropy of the GW signal across the sky is highly dependent on the underlying demographics and finiteness of the SMBBH population, rather than the actual positional distribution of sources. Current estimates place deviations from isotropy in the nanohertz GW sky below $\sim 10\%$ [56–59], making anisotropy a challenging yet rewarding scientific goal after the initial isotropic GWB detection by PTAs. The assumption of GWB isotropy leads to a simple analytic expression for the induced interpulsar timing correlations that is only dependent on the angular separation of the pulsars; this is known as the Hellings and Downs curve [60]. Correlations from anisotropic signatures can also be expressed analytically through a spherical harmonic decomposition of the angular power [56,57,61,62]. Current search strategies for anisotropy implement a prior condition that blocks off unphysical regions of the spherical harmonic coefficients that would otherwise imply negative angular power [57,63]. This prior reduces the efficiency of posterior sampling, and prevents gradient-based sampling methods from being used. More general methods that map the polarization content of the GWB using CMB techniques [62,64,65], or decompose the angular power distribution on eigen-skies of the PTA response map [66,67], are also possible but have not yet been implemented within realistic analysis studies.

The techniques developed in this paper represent a major step forward in mapping realistic gravitational-wave skies with PTAs (beyond those in Taylor *et al.* [63]), corresponding to methods that can be easily embedded in modern PTA pipelines, and that better account for anisotropy arising from bright single sources. Anisotropy searches should be superseded by joint searches for a stochastic GW background plus individual binary signals that can be resolved out of the

binary ensemble. These kinds of searches will employ transdimensional search strategies that use the data to not only constrain the properties, but also the number, of favored sources [43,44,68]. Sky localization of these individual binaries will initially be quite poor ($\sim \mathcal{O}(100)$ deg²) [50,55,69], necessitating known electromagnetic-counterpart candidates [70–77], or priors on likely host-galaxy properties to permit finer resolution [58,69,78].

After the initial detection of an isotropic GWB, and prior to the resolution of individual bright binaries within this background, these binaries will announce themselves merely as regions of excess angular power. Anisotropy searches only constrain the incoherent angular power distribution, so there is no Occam penalty for the many binary model parameters that otherwise appear in individual source searches; the drawback of course is that no binary characterization is possible. Probes of anisotropy form an important link in the chain of scientific milestones for PTAs, bridging the information gap between background detection and individual source resolution. We study this link here, developing fast and flexible mapping strategies for the angular power distribution of the GWB, and studying which of these is most apt for the GW sky from a SMBBH population. We employ Bayesian model selection to make data-driven decisions on the number of excess power regions that can be isolated, and forecast the signal-to-noise ratio at which anisotropy will be favored over isotropy.

This paper is laid out as follows. In Sec. II we introduce the formalism for computing the overlap reduction function between arrays of pulsars in the presence of an anisotropic stochastic GW background. We also discuss techniques to rapidly compute this under generic map parametrizations. In Sec. III we introduce our PTA likelihood for fitting simulated data to models, which is a simplified version of the full production-level model used in real PTA GW searches. In Sec. IV we discuss the two different spatial distributions of GW power that we consider here, corresponding to a region of excess power (a “hotspot”) and a map composed of many individual SMBBH signals drawn from a population synthesis simulation. We present our results in Sec. V, followed by concluding remarks in Sec. VI, and supplementary calculations in the Appendices.

II. ANISOTROPY

In PTA searches for a Gaussian, stationary stochastic GW background, the primary detection statistic is one that leverages two-point interpulsar correlated timing deviations. At the center of this statistic lies the overlap reduction function (ORF), which describes the degree with which GW-induced timing deviations are correlated between Earth-pulsar systems that vary in orientation with respect to angular power on the sky. The pieces needed to compute the ORF include the antenna response function of each Earth-pulsar system, along with a representation of the

underlying angular distribution of GW power. It is the latter dependence that allows us to probe angular structure in the nanohertz GW sky, in a bid to search for bright sources, clustered power regions, or statistical isotropy. In the following we describe different parametrized representations of this power, and test these with simulations to understand the most apt and statistically parsimonious representation of a given GW sky.

A. Antenna response

A GW in general relativity can be written as a sum over its two tensor transverse polarizations:

$$h_{\mu\nu}(t, \hat{k}) = e_{\mu\nu}^+(\hat{k})h_+(t, \hat{k}) + e_{\mu\nu}^\times(\hat{k})h_\times(t, \hat{k}), \quad (1)$$

where \hat{k} is the direction of GW propagation, $h_{+,\times}$ are the polarization amplitudes of the metric perturbation, and $e_{\mu\nu}^{+,\times}$ are the polarization basis tensors. The polarization basis tensors are constructed from a basis triad, such that

$$e_{\mu\nu}^+(\hat{k}) = \hat{m}_\mu \hat{m}_\nu - \hat{n}_\mu \hat{n}_\nu, \quad e_{\mu\nu}^\times(\hat{k}) = \hat{m}_\mu \hat{n}_\nu + \hat{n}_\mu \hat{m}_\nu, \quad (2)$$

$$\begin{aligned} \hat{k} &= -\hat{\Omega} = -(\sin \theta \cos \varphi)\hat{x} - (\sin \theta \sin \varphi)\hat{y} - (\cos \theta)\hat{z}, \\ \hat{m} &= -(\sin \varphi)\hat{x} + (\cos \varphi)\hat{y}, \\ \hat{n} &= -(\cos \theta \cos \varphi)\hat{x} - (\cos \theta \sin \varphi)\hat{y} - (\sin \theta)\hat{z}, \end{aligned} \quad (3)$$

where $\hat{\Omega}$ is the GW origin direction, and $\hat{x}, \hat{y}, \hat{z}$ are the Cartesian unit vectors. The equatorial declination and right ascension (δ, α) are related to the spherical polar coordinates by $\theta = \pi/2 - \delta$ and $\phi = \alpha$, where the north celestial pole is in the \hat{z} direction, and the vernal equinox is in the \hat{x} direction. From this point onwards we consider only $\hat{\Omega}$ and the origin location of GWs.

For a pulsar in direction \hat{p} , the antenna response function is the contraction of the polarization basis tensor with the impulse response function of an Earth-pulsar system to a GW, such that

$$\mathcal{F}^A(\hat{p}, \hat{\Omega}) \equiv \frac{1}{2} \frac{\hat{p}^\mu \hat{p}^\nu}{1 - \hat{\Omega} \cdot \hat{p}} e_{\mu\nu}^A(\hat{\Omega}), \quad (4)$$

which for each polarization is

$$\begin{aligned} \mathcal{F}^+(\hat{p}, \hat{\Omega}) &= \frac{1}{2} \frac{(\hat{m} \cdot \hat{p})^2 - (\hat{n} \cdot \hat{p})^2}{1 - \hat{\Omega} \cdot \hat{p}}, \\ \mathcal{F}^\times(\hat{p}, \hat{\Omega}) &= \frac{(\hat{m} \cdot \hat{p})(\hat{n} \cdot \hat{p})}{1 - \hat{\Omega} \cdot \hat{p}}. \end{aligned} \quad (5)$$

At a given frequency, the stochastic gravitational-wave background can be written as a superposition of plane waves, such that

$$\begin{aligned} h_{\mu\nu}(f, \mathbf{x}) &= \int_{S^2} d\hat{\Omega} \{h_+(\hat{\Omega}, f)e_{\mu\nu}^+(\hat{\Omega}) + h_\times(\hat{\Omega}, f)e_{\mu\nu}^\times(\hat{\Omega})\} e^{2\pi i f \hat{\Omega} \cdot \mathbf{x}/c}, \end{aligned} \quad (6)$$

where c is the speed of light, and we integrate over the whole sky S^2 . In pulsar timing, the integrated effect of the GWs on the pulsar times of arrival (TOAs) depends on the difference in the metric at the Earth (for convenience, the Solar System barycenter (SSB)) and at the position of the pulsar [79,80]. For a pulsar p in direction \hat{p} , the two contributions to the redshifted arrival rate of TOAs are known as the *Earth term*, s_p^E , and the *pulsar term*, s_p^P . These can be written in the frequency domain as (see Gair *et al.* [62] for full details)

$$\begin{aligned} s_p^E(f) &= \sum_{A=+,\times} \int_{S^2} d\hat{\Omega} \mathcal{F}^A(\hat{p}, \hat{\Omega}) h_A(\hat{\Omega}, f), \\ s_p^P(f) &= - \sum_{A=+,\times} \int_{S^2} d\hat{\Omega} \mathcal{F}^A(\hat{p}, \hat{\Omega}) h_A(\hat{\Omega}, f) e^{-2\pi i f L_p (1 - \hat{\Omega} \cdot \hat{p})/c}, \end{aligned} \quad (7)$$

where L_p is the distance to pulsar p , and the total signal is $s = s^E + s^P$. Appealing to stationarity of the GW background signal, the pulsar term is identical to the Earth term at each frequency except for a phase change, such that $h_{A,p}(\hat{\Omega}, f) = h_{A,e}(\hat{\Omega}, f) \equiv h_A(\hat{\Omega}, f)$. Due to the size of L_p , the pulsar term oscillates very rapidly across the sky, and much faster than our typical angular resolution. For all practical purposes in stochastic background analyses, we therefore regard the pulsar term phase as random.

Although it is possible to work out some of the integrals of Equation (7) analytically for various parametrized power representations on the sky [62], we can achieve complete generality by rephrasing the signal response as discretized sums over sky pixels using HEALPix.¹ With that, our Earth term GW sky $h_A(\hat{\Omega})$ now becomes $h_{(aA)}$, where the index a refers to the a th pixel in the sky, at location $\hat{\Omega}_a$. The antenna pattern functions $\mathcal{F}^A(\hat{p}, \hat{\Omega})$ become $F_{p(aA)}$, where p is the pulsar index as before. Henceforth, we use the following notation for indices: (a, b) refer to HEALPix pixel indices; (A, B) refer to GW polarization; and (i, j) refer to combined pixel polarizations, where $i = (aA)$. The matrix $F = \{F_{pi}\}$ is what we refer to as the signal response matrix, and the vector $\mathbf{h} = \{h_i\}$ is the full GW sky.² As for the GW signal, we split up the signal response matrix into an Earth

¹Hierarchical Equal Area isoLatitude Pixelation; <http://healpix.sourceforge.net>.

²Although we have $2N_{\text{pix}}$ degrees of freedom in \mathbf{h} , Gair *et al.* [62] showed that half of the GW sky is invisible to pulsar timing arrays, as all the curl modes have vanishing response.

term F^E and a pulsar term F^P . Since the extra phase of the pulsar term is a constant, we absorb it in the signal response matrix such that

$$F_{p(aA)}^E = \sqrt{\frac{3}{2}} \frac{\mathcal{F}^A(\hat{p}, \hat{\Omega}_a)}{N_{\text{pix}}} \quad (8)$$

$$F_{p(aA)}^P = F_{p(aA)}^E e^{-2\pi i f L_p(1 - \hat{\Omega}_a \cdot \hat{p})/c}, \quad (9)$$

where N_{pix} is the total number of sky pixels, and we have included a factor of $\sqrt{3/2}$ for a normalization convention that will become apparent in Sec. II C. We can thus write Eq. (7) in matrix notation:

$$\mathbf{s}(f) = \mathbf{F}\mathbf{h}(f), \quad (10)$$

with $\mathbf{F} = \mathbf{F}^E + \mathbf{F}^P$ as before. In this single frequency formalism, the size of \mathbf{s} is the number of pulsars N_{psr} , the size of \mathbf{h} is *twice* the number of pixels (to account for the two polarizations), and thus the size of \mathbf{F} is $(N_{\text{psr}} \times 2N_{\text{pix}})$. The array response matrix \mathbf{F} is a constant design matrix for all types of GWB anisotropy. By choosing a sufficient HEALPix resolution, the results from this discretization are indistinguishable from any analytic results. Unless otherwise explicitly stated, in the following we set the HEALPix $N_{\text{side}} = 32$, which gives $N_{\text{pix}} = 12288$. This is fine enough to provide us with fast and accurate numerical solutions.

B. Overlap reduction functions

For a wide-sense stationary Gaussian GW background, we can write

$$\langle h_{(aA)} h_{(bB)}^\dagger \rangle = \delta_{ab} \delta_{AB} \tilde{P}_{(aA)} = \delta_{ab} \delta_{AB} \frac{P_a}{2}, \quad (11)$$

where we denote ensemble averaging with $\langle \dots \rangle$, Hermitian transposition with \dagger , and $\delta_{(\dots)}$ is the Kronecker delta. The GW power at location $\hat{\Omega}_a$ in polarization A is $\tilde{P}_{(aA)}$. We also assume that $\tilde{P}_+ = \tilde{P}_\times = P/2$, where P is the total angular power summed over both polarizations. In all of the following the factor of $1/2$ has already been implicitly absorbed into the prefactor of Eq. (8).

The overlap reduction function Γ of the GW signal is defined as the reduction in signal correlation between two separated detectors [81], which in our case corresponds to different Earth-pulsar systems: $\Gamma_{op} = \langle s_p s_o^\dagger \rangle$. Using the signal response matrix formalism of Eq. (10), the ORF is trivially constructed from Eqs. (10) and (11):

$$\Gamma_{op} = \sum_i F_{oi} P_i F_{pi}^\dagger. \quad (12)$$

Note that the Earth term and pulsar term components are numerically orthogonal to each other due to the rapid

oscillation of the pulsar term across the sky with respect to the Earth term. For the same reason, the pulsar-term contribution to the overlap reduction function is effectively diagonal, meaning we can simply double the diagonal Earth-term ORF components to successfully replicate the full ORF. As such, we have

$$\begin{aligned} F_{oi} P_i F_{pi}^\dagger &= F_{oi}^E P_i F_{pi}^{E\dagger} + \delta_{oi} F_{oi}^P P_i F_{pi}^{P\dagger} \\ &= (1 + \delta_{op}) \times F_{oi}^E P_i F_{pi}^{E\dagger}, \end{aligned} \quad (13)$$

which we can write in compact matrix form as

$$\begin{aligned} \mathbf{\Gamma}^E &= \mathbf{F}^E \cdot \mathbf{P} \cdot \mathbf{F}^{E,T}, \\ \Gamma_{pp} &= 2\mathbf{\Gamma}_{pp}^E, \quad \Gamma_{op} = \mathbf{\Gamma}_{op}^E, \end{aligned} \quad (14)$$

where $\mathbf{\Gamma}, \mathbf{\Gamma}^E$ is the $(N_{\text{psr}} \times N_{\text{psr}})$ array ORF matrix for the full signal and Earth-term-only signal, respectively. We also note that normal transposition has replaced Hermitian transposition since \mathbf{F}^E is real. The diagonal $(2N_{\text{pix}} \times 2N_{\text{pix}})$ matrix \mathbf{P} contains the array of power values in each sky pixel along the diagonal, with $+$, \times polarizations as pixel neighbors.

C. The Hellings and Downs curve

As a sanity check, we recover the ORF for an isotropic background (known as the *Hellings and Downs curve* [60]), by modeling $\mathbf{P} = \mathbb{I}_{2N_{\text{pix}}}$. For an arbitrary array of pulsars we can construct the signal response matrix and carry out the matrix operations in Eq. (14), which correspond to a double integral over the sky. When integrating analytically, the resulting ORF expression is a function only of the angular separation of pulsars rather than their absolute position. However, tackling this integral analytically requires a combination of coordinate transformations and contour integration [56,62,79], all of which has been avoided by casting the problem into matrix algebra.

For an array of 100 pulsars drawn from an isotropic distribution on the sky, we compare our numerical results to the following analytic expression [which is normalized by convention to equal 1 for $o = p$, thus explaining the $\sqrt{3/2}$ factor adopted in Eq. (8)]:

$$\Gamma(\xi_{op}) = \frac{3}{2} x \ln(x) - \frac{x}{4} + \frac{1}{2} + \frac{\delta_{op}}{2}, \quad x = \frac{1 - \cos \xi_{op}}{2}, \quad (15)$$

where δ_{op} accounts for the pulsar term contribution. This comparison is shown in Fig. 1, where in the top panel the numerical ORF values are offset from the analytic ORF values by $+0.2$ for ease of viewing. The lower panel shows that the absolute difference between the numerical and analytic values remains $\lesssim 10^{-4}$ over the entire angular separation range. Our numerical ORF formalism is fast, flexible, and accurate.

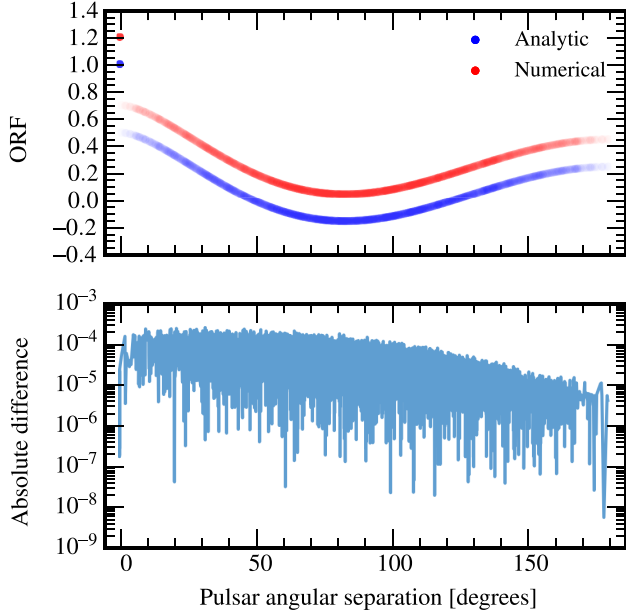


FIG. 1. Comparison of analytic and numerical *Hellings and Downs* curve computation for a 100-pulsar array. The analytic calculation uses Eq. (15), while the numerical calculation assumes an isotropic power map within Eq. (14). In the top panel, the numerical calculation is offset from the analytic one by +0.2 for ease of viewing. The lower panel shows the absolute difference between the numerical and analytic calculations.

D. GW map parametrizations

In a pixelated power representation such as ours, the most obvious parametrized power representation is to search for the P_i values individually at a predetermined pixel resolution. This is an inelegant and inefficient approach though, since the number of parameters in the model gets prohibitively large, and PTAs typically do not have a high response resolution. In the following we list several parametrizations that are appropriate for different signal scenarios. One of these has appeared already in the literature (the spherical-harmonic model) while the rest are new to this paper. While we discuss these parametrizations separately, the models can be linearly combined where necessary to fit broad anisotropic features (spherical-harmonic or disk) and localized features (point sources).

1. Spherical harmonic modeling of power

The spherical-harmonic basis is an obvious choice to decompose a scalar field on the 2-sphere, since the lowest order basis function describes isotropy, and higher functions add anisotropic power at smaller angular scales. Spherical-harmonic power reconstructions were first proposed in this context for LIGO sky mapping [82], but have recently received significant attention in the PTA literature [56,57,62,63], where the power is decomposed as

$$P(\hat{\Omega}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\hat{\Omega}), \quad (16)$$

where Y_{lm} are the real-valued spherical harmonics, and c_{lm} are the spherical harmonic coefficients of $P(\hat{\Omega})$. In this parametrization, an isotropic sky is described by $\{c_{00} = \sqrt{4\pi}\}$, while a maximal dipole in the $+\hat{z}$ direction is described by $\{c_{00} = \sqrt{4\pi}, c_{10} = \sqrt{4\pi/3}\}$. This linear parametrization of the power in terms of the c_{lm} coefficients leads to a similar expression for the ORF:

$$\Gamma_{op} = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{lm} \Gamma_{(lm)(op)}, \quad (17)$$

where $\Gamma_{(lm)(op)}$ are the anisotropy ORF components for anisotropy components (l, m) between pulsars p and o , which can be precomputed for a specific PTA from just the pulsar sky locations.

Although Mingarelli *et al.* [56] and Gair *et al.* [62] have given detailed analytic formulas to evaluate the $\Gamma_{(lm)(op)}$ for any (lm) , in practice it is equally efficient to use linear algebra to compute these quantities numerically using Eq. (12):

$$\Gamma_{(lm)(op)} = \sum_i c_{lm} Y_{(lm)i} F_{oi} F_{pi}^\dagger, \quad (18)$$

where $Y_{(lm)i}$ is the value of $Y_{(lm)}$ evaluated at sky location $\hat{\Omega}_i$. In this formalism, the model parameters are the coefficients c_{lm} . Normalization is ensured by fixing $c_{00} = \sqrt{4\pi}$ since the $l \geq 1$ spherical harmonics are orthonormal to the constant $l = 0$ function (where $Y_{00} = 1/\sqrt{4\pi}$), and thus themselves integrate to zero across the sphere. Since c_{00} is fixed, variation of the monopolar level of the GWB is absorbed into the strain amplitude parameter.

For a given set of $\{c_{lm}\}$ to represent a physical power map, the resulting power has to be non-negative at every point on the sky. In practice this is performed by evaluating Eq. (16) on a grid of test sky positions for a set of proposed $\{c_{lm}\}$ [57]. In this paper we sample uniformly in $c_{lm} \in U[-5, 5]$ where $l \geq 1$, and evaluate the *physical prior* on a HEALPix $N_{\text{side}} = 8$ sky, which gives 768 test positions. In a model with l_{max} as the multipole cutoff, the number of anisotropy parameters is $(l_{\text{max}} + 1)^2 - 1$. In Sec. IV we study how sky reconstruction and the Bayes factor for anisotropy varies with the choice of l_{max} , up to $l_{\text{max}} = 6$.

2. Spherical-harmonic modeling of square root of power

While the spherical harmonics are a natural basis with which to decompose the GWB power, the practical implementation has required us to introduce a nonanalytic prior on the c_{lm} coefficients. We can bypass this rejection prior by instead parametrizing the *square root of power*, which is

allowed to be negative, and will by construction always produce a positive power distribution. The benefits are twofold—(i) the prior is now completely analytic; and (ii) sampling performance is improved through better mixing, shorter burn-in, and the potential to propose draws from the prior volume during sampling.³

We decompose the square root of power in exactly the same way as for the power:

$$P(\hat{\Omega})^{1/2} = \sum_{L=0}^{\infty} \sum_{M=-L}^L a_{LM} Y_{LM}(\hat{\Omega}), \quad (19)$$

where Y_{LM} are the real-valued spherical harmonics, and a_{LM} are search coefficients. We have capitalized the component labels $\{LM\}$ to be distinct from the usual $\{lm\}$ components of the GWB power.

There is not a direct mapping between structure in l of the power and structure in L of the root power, e.g., a dipole in power will be described by $l_{\max} = 1$, but in principle requires contributions in the root power from $L > 1$. We show this analytically in Appendix A, where the power coefficient c_{lm} is decomposed as an infinite summation of a_{LM} terms. However, we also show that adequate reconstruction of a map with $l \leq l_{\max}$, can be achieved with $L \leq L_{\max} = l_{\max}$.⁴

3. Point-source anisotropy

The spherical-harmonic model builds towards smaller-scale anisotropic structure through successively higher multipoles. Thus if the GW sky is predominantly isotropic but with one bright SMBBH sticking prominently above the background level, we will need a high l_{\max} (and consequently a large number of parameters) to localize it. Instead, we can model an isolated point of GWB power with a 2-sphere delta function, such that $P(\hat{\Omega}) = P' \delta^2(\hat{\Omega}, \hat{\Omega}')$. The corresponding physical picture is of a *stochastic point source* [84–86].

We implement this point source of power by lighting up a single pixel in our modeled sky, and keeping all other pixels dark. The only search parameters are for the sky location of the lit pixel, which can be referred to through pixel indices. The power in this pixel is set to N_{pix} to satisfy normalization, and the ORF is computed using Eq. (14). The source sky location is searched over with a uniform prior on the sky, i.e., $\alpha \in U[0, 2\pi]$, $\cos(\pi/2 - \delta) \in U[-1, 1]$. We can combine these point sources with different amplitude weightings (sometimes in addition to an isotropic background) to model

³The ability to implement draws from the prior aided particularly in sampling posterior multimodalities that derived from a given $P(\hat{\Omega})$ having two possible differing-sign solutions for $P(\hat{\Omega})^{1/2}$.

⁴Recently, a similar method has been proposed in the context of ground-based GW constraints on the anisotropy of BBH merger events [83].

whatever small-scale angular features are supported by the data.

In Sec. IV we study how sky reconstruction and the Bayes factor for anisotropy varies as we model the sky with different numbers of point sources, from $N_{\text{point}} = 1, \dots, 5$. The scenario of multiple modeled point sources stacking on top of each other during sampling is avoided through the natural parsimony of Bayesian model selection. We help this along during the posterior sampling by implementing a prior that tests and rejects any new proposed point-source location that would be cast to the same pixel index as another in an $N_{\text{side}} = 8$ sky. We note that combining this point-source model of anisotropy with a frequency-dependent amplitude parameter permits a power-based search for bright single GW sources that can complement existing continuous-wave and burst methods.

4. Disk anisotropy

If the angular power distribution reflects the local clustering of host galaxies to produce an extended *hotspot* region on the sky, then both the spherical-harmonic and point-source models have drawbacks. We can combine the ability of the spherical-harmonic approach to model broad anisotropic features with the ability of the point-source model to localize bright regions of power. This hybrid is a *disk* of GW power with a modeled sky location and angular radius. The power is constant inside the disk perimeter and zero outside. This is implemented using the HEALPix `query_disc` function, which returns an array of pixel indices lying within a disk of supplied sky location and angular radius. The disk model has three search parameters: two for sky location (with a uniform prior on the sky) and one for the angular radius (with prior $\theta_r \in U[0, \pi]$ to avoid the disk wrapping around the sphere and overlapping with itself).

III. LIKELIHOOD

We significantly simplify the pulsar-timing likelihood to permit fast analysis of datasets with a wide range of different mapping parametrizations. We are only interested in the comparative modeling and detection efficacy of different GWB map parametrizations, and so ignore the pulsar ephemeris model. The dominant effect of this ephemeris model on signal inference is to remove sensitivity to lower frequencies due to the subtraction of a quadratic function (which models the pulsar spin-down). It also diminishes sensitivity at sampling frequencies of 1 and 2 yr^{-1} due to the fitting of astrometric parameters. Likewise, sophisticated noise modeling is essential to properly isolate GW signals from noise processes, but we do not include processes beyond radiometer noise here. Nevertheless, we stay true to the spirit of the likelihood as it is used in real searches (e.g., [17,87]), while stripping out all the model contributions that are irrelevant to this paper.

A. Assumptions

In our simplified model, we have N_p pulsars with N_{obs} observations each. For pulsar p the observations $\delta \mathbf{t}_p$ consist of two components:

$$\delta \mathbf{t}_p = \mathbf{s}_p + \mathbf{n}_p, \quad (20)$$

where \mathbf{s}_p is the GW signal and \mathbf{n}_p is the noise. Given our focus on GWB map parametrizations, we assume a simple Gaussian white-noise model in our simulations: $\langle n_{pc} n_{od} \rangle = \sigma_{p,c}^2 \delta_{op} \delta_{cd} = N_{p,cd} \delta_{op} \delta_{cd}$, where $\{p, o\}$ index pulsars, $\{c, d\}$ index observations, and $\sigma_{p,c}$ is the uncertainty in a single observation. This uncertainty is usually referred to as the radiometer noise [87], which we here assume to be a known homoscedastic quantity for each pulsar. The $\langle \dots \rangle$ indicates ensemble averaging (equal to the time average here since we assume ergodicity).

We model the GW signal, \mathbf{s}_a , as a zero-mean, wide-sense stationary stochastic background with the covariance between observation c in pulsar p and observation d in pulsar o parametrized as

$$\langle s_{pc} s_{od} \rangle = \Gamma_{po} C(|t_c - t_d|), \quad (21)$$

where Γ_{po} is the ORF between these two pulsars (i.e., the spatial correlation), and $C(|t_c - t_d|)$ is the time-domain correlation induced by the GWB.

For the sake of simplicity we further assume that the GWB signal has been whitened, such that $C(|t_c - t_d|) = A^2 \delta_{cd}$. This assumption in no way affects the interpretation of our results in this paper, since we are only interested in the relative efficacy of power-map parametrizations, and not spectral parametrizations. The true GWB from a population of SMBHB will induce long-timescale stochastic deviations in the TOAs, with an $f^{-13/3}$ power-law spectral density [34–36, 88, 89] that corresponds to a dense time-domain covariance matrix. This means that in real searches most information about the GWB is derived from the lowest few sampling frequencies in the TOA data. In such real analyses the dense time-domain covariance is modeled through rank-reduced means [90–92], or through data whitening based on estimates of the covariance [93]. Assuming a white GWB spectrum here allows us to rapidly compute the PTA likelihood. However, our modeling of anisotropy and the PTA overlap reduction function requires very little effort to incorporate into real production-level pipelines that deal with multiple GW frequencies and unevenly sampled time-series data, which we discuss in Sec. VI.

B. Marginalized likelihood

We marginalize over the low-level linear parameters \mathbf{s}_p and \mathbf{n}_p with Gaussian priors [17, 87, 90, 91]. This gives a likelihood that depends only on model hyperparameters: the

GWB amplitude A ; the TOA uncertainties $\{\sigma_p\}$ (assumed to be the same for all observations and all pulsars, with a delta-function prior set at the measured uncertainties); and the GWB map parameters $\{c_p\}$. The log likelihood $\ln \mathcal{L} = \ln p(\delta \mathbf{t} | A, \sigma_p, c_p)$ for all pulsars is then given by

$$\ln \mathcal{L} = \sum_{c=1}^{N_{\text{obs}}} \left\{ -\frac{1}{2} \delta \mathbf{t}_c^T \Sigma_c^{-1} \delta \mathbf{t}_c - \frac{1}{2} \ln \det 2\pi \Sigma_c \right\}, \quad (22)$$

where $\delta \mathbf{t}_c$ is a vector of size N_p consisting of the c th observation for each pulsar, and

$$\Sigma_c = A^2 \Gamma + \mathbf{N}_c, \quad (23)$$

where Σ_c is the $(N_p \times N_p)$ inter-pulsar correlation matrix of the c th observation, Γ is the $(N_p \times N_p)$ dense ORF matrix of the c th observation (we assume that it is the same for all observations), and \mathbf{N}_c is the $(N_p \times N_p)$ diagonal noise covariance matrix of the c th observation. The latter noise matrix is assumed to be the same for all pulsars and all observations, and remains fixed in our sampling. In production-level PTA analysis pipelines, the covariance matrices of long-timescale stochastic processes (like the GWB) are modeled on rank-reduced bases [90–92], e.g., a Fourier basis. It is thus straightforward to generalize our ORF modeling to permit different ORFs at each frequency. Additionally, generalizing Eq. (22) to account for data taken at different times in each pulsar is most easily achieved by concatenating all observations into a single data vector, with \mathbf{N} still diagonal over the entire PTA, and $A^2 \Gamma$ now equal to the full-array version of Eq. (21).

We apply a uniform prior in $\log_{10}(A/\text{secs})$ over a sampling range $[-10, -3]$. All remaining variables in the model are used to parametrize the GWB map, except in the case of multiple point sources, where each component, k , receives a separate amplitude parameter A_k to permit different source weightings. The priors on $\{c_p\}$ are listed in Sec. II for each parametrization. The $A^2 \Gamma$ term describes the temporal and spatial correlations induced by the GWB, which can be expressed through linear combinations of anisotropic features using the parametrized models in this paper, e.g., an isotropic background with multiple bright binaries has $A^2 \Gamma = A_{\text{iso}}^2 \Gamma_{\text{iso}} + \sum_{k=1}^{N_{\text{comp}}} A_k^2 \Gamma_k$.

IV. SIMULATIONS

A. Tested GW sky maps

We test the relative efficacy of our mapping parametrizations on two qualitatively different types of GW skies, shown in Fig. 2 with their associated angular power spectrum. These are the following:

- (i) *Hotspot*: an extended hotspot of power that dominates the entire GW sky. This is an edge case that in less extreme scenarios may represent a cluster of galaxies

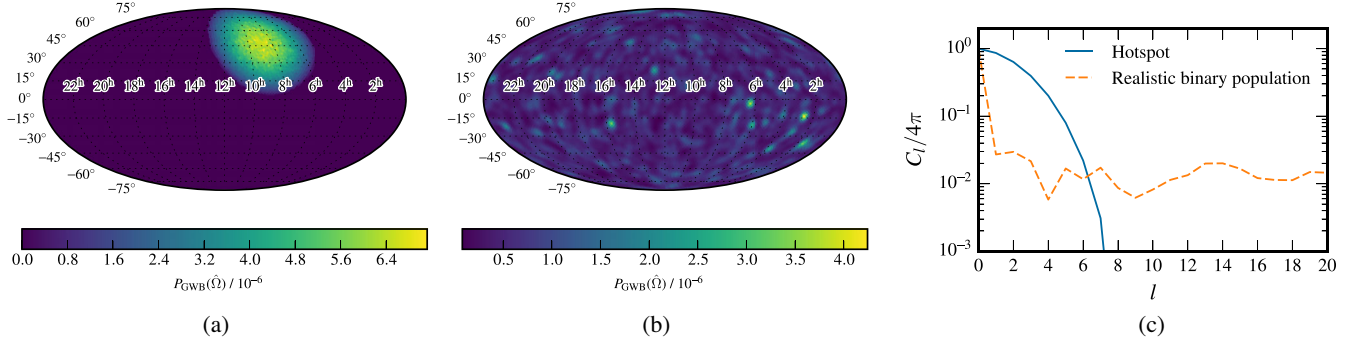


FIG. 2. The GW skies that we test our models against in this paper. The map in (a) represents a hotspot region of excess GW power, while the map in (b) is a smoothed map generated from a synthesized population of SMBBHs. The angular power spectrum of these maps is shown in (c).

hosting many merging SMBBHs, or a local Universe inhomogeneity in the SMBBH distribution. This map was constructed by performing an $l \leq 5$ decomposition on an isotropic map with a large value added to one pixel. The extended region of power was excised and added to an otherwise empty map.

- (ii) *SMBBH population*: a more realistically diffuse GW sky created from the output of a SMBBH population synthesis simulation constructed with the technique described in [48]. The SMBBH population is constructed empirically based on a combination of observed quantities and theoretical input. In short, a galaxy merger rate is constructed by multiplying the galaxy mass function to the observed galaxy pair fraction and by dividing it for the characteristic merger timescale, which is estimated from numerical simulations. Galaxies are then populated with SMBHs according to scaling relations drawn from the literature. The model adopted here combined the galaxy mass function from [94], pair fraction from [95], merger timescale from [96], and SMBH-bulge mass relation from [97], to produce a catalog of SMBBHs with chirp masses, redshifts, and orbital frequencies from which the GW strain of each (assumed circular) binary was computed. By selecting those systems with observed GW frequencies between 0.1 and 0.2 yr^{-1} , only binaries within the lowest frequency resolution bin of a PTA with ten-year timing baseline were collected to create the sky map. The angular positions were assigned to be statistically isotropic, but the strain distribution is dominated by only a small number of GW-bright systems, producing anisotropy in the angular power distribution. Most SMBBH systems in this map constitute a background of sources that are not individually resolvable, with only a handful of systems emitting signals that resound prominently above this. The displayed map has been smoothed with a Gaussian kernel to allow structure to be seen; the true map looks mostly isotropic with only a few dominant pixels due to bright systems.

In the following we describe the format of the tested datasets that contain the imprint of these GW power distributions, and the suite of signal injections.

B. Simulation setup

Our interest lies in the most apt model for GWB anisotropy rather than finding the optimal geometrical properties of a PTA. Hence, we create a statistically isotropic array of 100 pulsars on the sky, as may be operational in the mid- to late-2020s when probes of signal anisotropy will be possible. Current arrays use ~ 40 – 50 pulsars in flagship GW searches [98,99] (although NANOGrav now times as many as 79 pulsars [100]), so forecasting an additional ~ 6 pulsars per year to be added to the IPTA is reasonable [101]. As discussed in Sec. III A, we do not concern ourselves with the pulsar timing model nor differing noise characteristics. The pulsars only differ through their geometrical response to a GW signal, as characterized by their position on the sky. Each pulsar has 100 observations that are evenly sampled and concurrent across the array, with homoscedastic white-Gaussian noise of standard deviation 100 ns.

For reasons discussed in Sec. III A, the GWB is modeled as a whitened Gaussian process of standard deviation A (which is varied in the injections) that is spatially correlated across the array. This spatial correlation depends on the ORF and thus the angular power distribution of the GW sky. Since the GWB is whitened, a single PTA dataset with 100 observations per pulsar can be considered as a concatenation of (for example) 20 separate realizations of a signal with 5 degrees of freedom (i.e., informative frequencies), which is representative of signals that are likely to be recovered in the near future. Hence we create many datasets with different values of A for each sky, but the random seed used to produce the signal realizations is kept fixed, as is the seed for the noise injection, which is fixed at an independent value for each pulsar. The hotspot datasets have signals injected on an evenly spaced grid of 50 values of A in the range $[5, 50]$ ns. The realistic SMBBH

population datasets have signals injected on an evenly spaced grid of 100 values of A in the range $[10, 500]$ ns.

V. RESULTS

An isotropic background is a convenient baseline model with which we compare the efficacy of different sky-map parametrizations. Indeed, for our chosen PTA configuration we can compute the normalized geometrical match between the ORF of our two different skies with the corresponding isotropic ORF. This represents the theoretical limiting match between the two ORFs in the infinite SNR limit, taking the form [102,103]

$$M = \sum_{p,o \neq p} \frac{\Gamma_{po}^{\text{sky}} \Gamma_{po}^{\text{iso}}}{(\Gamma_{po}^{\text{sky}} \Gamma_{po}^{\text{sky}})^{1/2} (\Gamma_{po}^{\text{iso}} \Gamma_{po}^{\text{iso}})^{1/2}}, \quad (24)$$

and has values of $M = 0.78$ for the hotspot sky and $M = 0.98$ for the SMBBH population sky. The signal correlations seen by our PTA are not so different from those induced by an isotropic GWB, even in the case of the hotspot sky.

The aforementioned statistic only accounts for the geometrical match of signal correlations, but the noise properties of the pulsars and the presence of the GW signal in the arrival time deviations should also be accounted for. In fact, introducing signal and noise weightings takes us to the signal-to-noise ratio (SNR). The general form of the expected value of this SNR is [79,104]

$$\langle \rho \rangle = \left(2T \sum_{p,o \neq p} \Gamma_{po}^{\text{true}} \times \Gamma_{po}^{\text{template}} \int_{f_l}^{f_h} df \frac{P_g^2(f)}{P_p(f) P_o(f)} \right)^{1/2}, \quad (25)$$

where T is the timing baseline, $P_g(f)$ is the spectrum of GWB-induced timing residuals, and $P_{p,o}(f)$ is the spectrum of noise-induced residuals in each pulsar. If radiometer noise is the only source of intrinsic per-pulsar noise, then $P_p(f) = 2\sigma_p^2 \Delta t_p + \Gamma_{pp} P_g(f)$, where σ is the rms uncertainty in the timing residuals and Δt is the effective timing cadence. Equation (25) accounts for the possibility of a template ORF being different from that of the true data. If the template ORF equals the true ORF, then we call $\langle \rho \rangle$ the *optimal* SNR; otherwise if we adopt an isotropic template this is the *suboptimal* SNR. We will now drop the expectation brackets for ease of readability. The ratio of the suboptimal SNR to the optimal SNR depends mildly on the strength of the GWB signal; in our simulations, the hotspot case has $(\rho_{\text{subopt}}/\rho_{\text{opt}}) \sim 0.59\text{--}0.62$, while for the SMBBH population $(\rho_{\text{subopt}}/\rho_{\text{opt}}) \sim 0.96\text{--}0.97$.

A. Hotspot sky tests

In order to guide our intuition in later cases, we use the hotspot sky for a first round of model selection with several

of our sky parametrizations. Since the map was initially constructed from an $l_{\text{max}} = 5$ restricted multipole decomposition of a single-pixel sky, we perform a spherical harmonic reconstruction with $l_{\text{max}} = 5$. We also test a disk and point-source model; in both of these cases, the angular power is zero except inside the pixel region of the disk/point. We do not test the $P(\hat{\Omega})^{1/2}$ spherical-harmonic decomposition on this hotspot sky, instead deferring this to the SMBBH sky where its full efficacy can be shown. These hotspot sky tests are not meant to be an exhaustive study of all methods, instead acting as a guide for our intuition and a test of our sky-mapping framework. We perform simultaneous Bayesian parameter estimation and model selection using the product-space sampling approach outlined in Appendix B. The relevant Bayes factor that we track is between a given anisotropy model and the baseline assumption of an isotropic sky.

Figure 3 shows that at low SNR the data is initially uninformative and unable to arbitrate any deviation from isotropy, giving $\mathcal{B} \sim 1$. As the SNR grows, all models except the point source have monotonically increasing Bayes factors. The initial tendency for the data to favor an isotropic sky over a point source is due to the extreme spatial compactness demanded by this model, with all power in one pixel. The data initially prefer isotropy, but as the SNR grows so does the data's tendency to favor this spatially compact description to explain the hotspot of power. As one would expect for this sky, a disk of power is the favored model overall, striking a balance between appropriateness to the power distribution and model parsimony. The results in Fig. 3 can be compared with those of Hotinli *et al.* [65], where a localized GW beam accounting for more than 50% of the total GW energy density can be detected at 3-sigma when the SNR ≥ 10 .

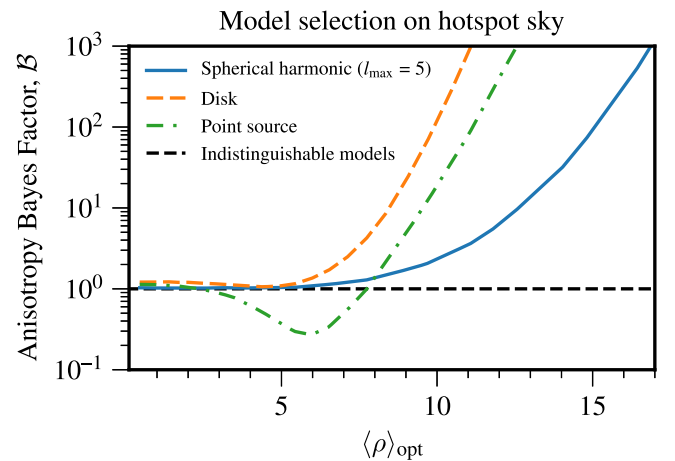


FIG. 3. The Bayes factor for anisotropy versus isotropy of the hotspot GW sky under different model parametrizations is shown here. The most apt model is clearly the disk and point-source descriptions, although the point source suffers an initial period of being disfavored.

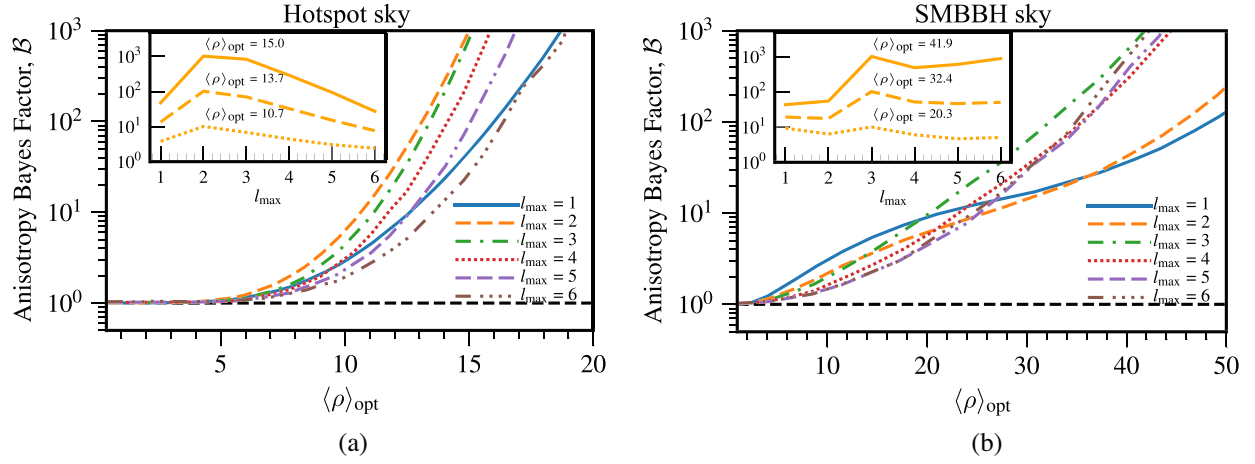


FIG. 4. Comparison of anisotropy-vs-isotropy Bayes factor growth with SNR for different choices of l_{max} in spherical-harmonic power modeling of (a) a hotspot sky, and (b) a realistic SMBBH sky. Inset figures show vertical slices through the curves at different SNR values.

While the spherical harmonic model seems to perform the worst overall in Fig. 3, it has been artificially restricted here to $l_{\text{max}} = 5$. We now relax this, showing in Fig. 4(a) the Bayes factor growth as a function of SNR and different l_{max} . The inset figure slices through the different l_{max} curves at fixed SNR, showing how the Bayes factor evolves with l_{max} under fixed signal strengths. An $l_{\text{max}} = 2$ model is favored at all SNRs, although there is a hint that $l_{\text{max}} = 3$ may be equally favored at higher SNRs, where more informative data counteracts the Occam penalty brought by more model parameters.

Finally, in the left column of Fig. 6(a) we show the power reconstructions of all these different parametrizations of the hotspot sky, including the favored $l_{\text{max}} = 2$ spherical harmonic model. The injected map is shown at the top. The map reconstructions are averaged over the relevant model's posterior probability distribution. In all cases the

region of hotspot power has clearly been identified and localized.

B. SMBBH sky tests

A sky map produced by a population of SMBBHs is more challenging than the simple hotspot sky, as there is no obvious model that will *a priori* be favored. The natural first test is between different l_{max} spherical harmonic recoveries, for which we show the Bayes factor growth versus SNR in Fig. 4(b). Note the SNR scale; arbitrating between anisotropic and isotropic power maps of SMBBHs may require very informative data. At low SNR the model with the fewest parameters is favored, corresponding to $l_{\text{max}} = 1$. As the data becomes more informative at SNRs ranging from 20–40, an $l_{\text{max}} = 3$ model is favored. The inset figure hints at higher multipolar models becoming less disfavored at higher SNRs, where the data is better able to

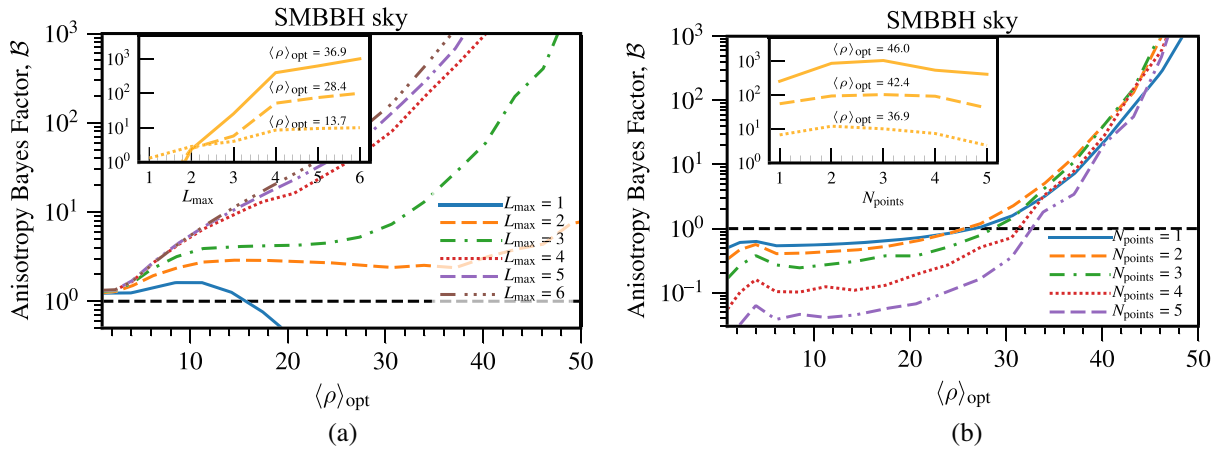


FIG. 5. Comparison of anisotropy-vs-isotropy Bayes factor growth with SNR in a SMBBH sky for different choices of (a) L_{max} in spherical-harmonic modeling of $P(\hat{\Omega})^{1/2}$, and (b) N_{points} in a multiple point-source model. Inset figures show vertical slices through the curves at different SNR values.

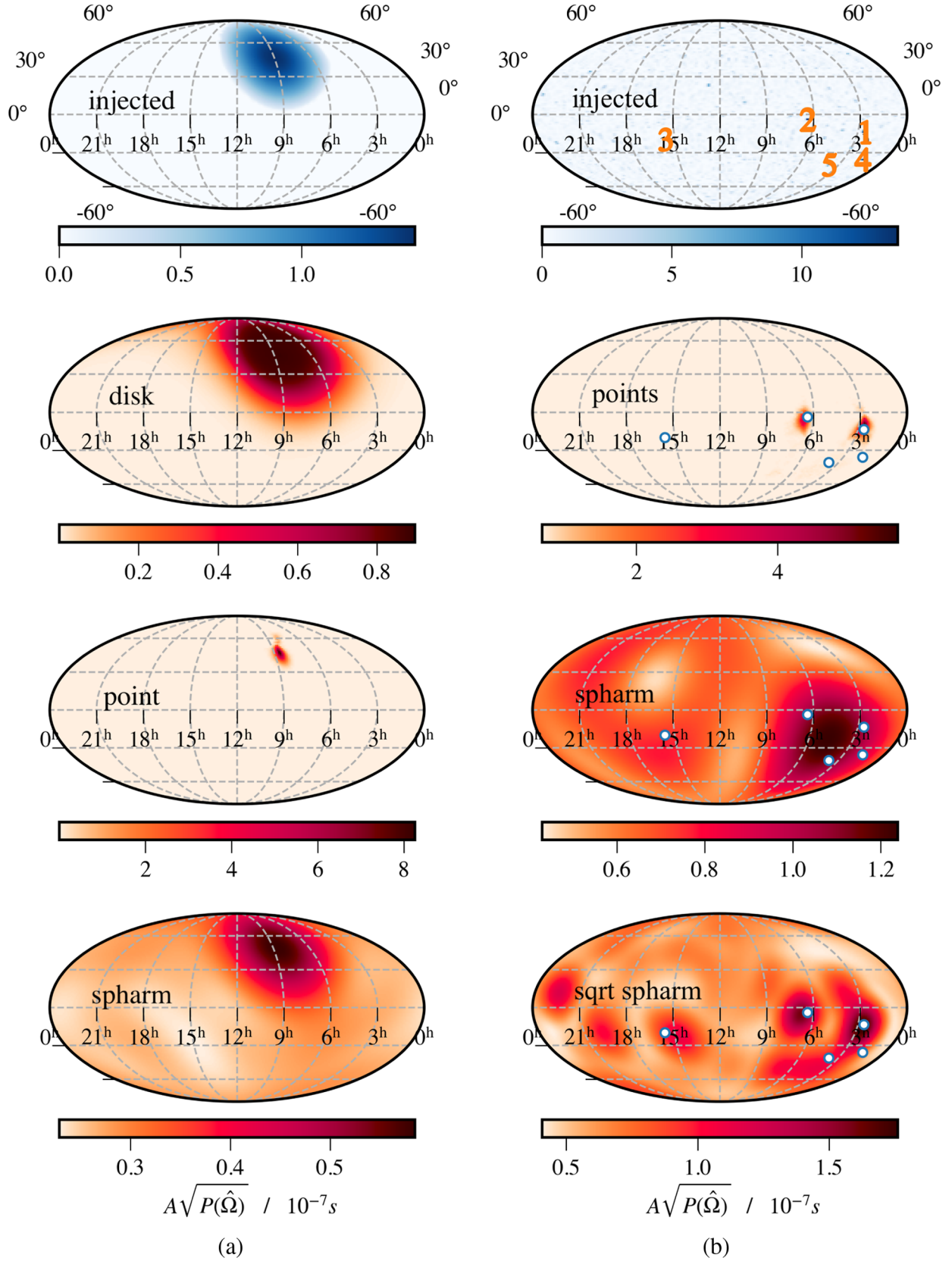


FIG. 6. Posterior averaged sky maps for different parametrized mapping models tested against (a) a hotspot sky, and (b) a realistic SMBBH sky. The injected maps are shown as the top two panels in blue, with all reconstructions shown in red. For the SMBBH sky, the top five ranked loudest GW binaries are shown as numbered locations in the injected map and blue-edge white circles in the recovered maps.

arbitrate deviations from isotropy and thus constrain smaller angular-scale features.

We next reconstruct the sky by modeling the square root of the power map with spherical harmonics. The benefits of this approach with respect to straightforward spherical-harmonic power modeling were discussed in Sec. IID. The results are shown in Fig. 5(a), where we see that $L_{\max} = 1$ is dramatically disfavored; the squaring of the map to return to $P(\hat{\Omega})$ produces features that are markedly different from isotropy and so are disfavored even at low SNR. However, increasing the L_{\max} has a remarkable impact on the Bayes factor growth, with the data preferring ever larger L_{\max} at all tested SNRs. The interpretation here is subtle, and related to how even a low- l power anisotropy requires contributions from $L > l$ multipoles in the $P(\hat{\Omega})^{1/2}$ model (see Appendix A). Despite favoring ever higher L_{\max} , there is clearly no Occam penalty associated with the accompanying increase in the parameter volume. In fact, $L_{\max} = 6$ reaches $\mathcal{B} = 100$ at $\text{SNR} = 28$, while $l_{\max} = 3$ passes the same threshold later at $\text{SNR} = 32$.

Finally, we employ the most physically motivated model for this kind of sky from the standpoint of its composition: point sources. The implementation differs from the previous hotspot-sky tests in two key ways: (1) as mentioned in Sec. IID, we allow for multiple point sources across the sky, each with its own amplitude parameter; (2) these point sources sit atop an isotropic GWB signal to emulate a realistic search scenario for multiple bright binaries resounding above a background. In Fig. 5(b) we can see that any number of points is initially disfavored, with the data preferring an isotropic model below $\text{SNR} \sim 25$. It is clear that the data lacks enough information to discriminate individual sources within the background at low SNR. Given that these point sources are being modeled on top of an isotropic signal, at low SNR the additional parameters of each source are simply incurring an Occam penalty. The penalty compounds for each additional modeled point source. It is only above $\text{SNR} \sim 25$ that the data can discriminate individual sources, and by $\text{SNR} \sim 35$ even preferring an Iso + 5pts model. However, this model still performs worse than the root-power spherical-harmonic model, with Iso + 2pts reaching $\mathcal{B} = 100$ at $\text{SNR} = 42$. Note that all of these SNR values at which various odds ratio thresholds are surpassed are specific to this SMBBH population realization. Hence while they are instructive and indicative of general trends, the SNR values at which anisotropy is favored may in fact be much lower if a small number of binaries swamp the entire GW sky at a given frequency.

The posterior-averaged power reconstructions of all these different parametrizations of the SMBBH population are shown in Fig. 6(b). The injected distribution is shown as the top map. The five binaries that are loudest in GWs are shown as numbered locations in the injected map, and blue-edged white circles in the reconstructed maps. Each

parametrization localizes important regions of power to varying degrees, with the point-source and sqrt-spharm models best able to discriminate individual sources. In the latter case this is not obvious from the parameter posteriors, requiring map recovery to really visualize the strength of this model.

VI. CONCLUSIONS

We have presented a general-purpose scheme to rapidly compute the interpulsar correlation signatures of anisotropic GW skies. This scheme is intended to be fast and flexible for embedding within a Bayesian analysis pipeline, and we have studied this use case here. We considered several different parametrized representations of anisotropic power; namely (i) a spherical-harmonic power decomposition, (ii) a spherical-harmonic decomposition of the square-root power, (iii) lighting up multiple pixels atop an isotropic background, (iv) a disk of power to model an extended region of GW emission. Our pipeline was tested against two very different power distributions, corresponding to an obvious region of excess power (a hotspot), and a realistic GW sky generated by a population of SMBBHs. A key component of our Bayesian modeling was model selection, to not only determine when the data favored anisotropy over isotropy, but also to optimize various model hyperparameters, e.g., the maximum multipole of spherical-harmonic modeling, and the number of pixels that were lit up to represent single sources.

For the hotspot sky, we found that the most apt model from a Bayesian perspective was a disk of power, with even a single point source being favored over spherical-harmonic models with $l_{\max} < 6$. This happens because Bayesian model selection will favor the most parsimonious model that can represent the information in the data: a disk or single pixel can do so here with only a few parameters, whereas spherical harmonics require $(l_{\max} + 1)^2$ parameters. For the SMBBH sky, our analysis marks the first time that robust Bayesian PTA techniques have been brought to bear on such a realistic GW sky. We found that the model that performed best was the spherical-harmonic modeling of square root of the power distribution, which avoided the restrictive prior condition of the usual spherical-harmonic power modeling, and which did not seem to suffer any Occam penalties through the inclusion of higher multipoles. Nevertheless, the model that lit up multiple pixels atop an isotropic GW background provides the most physically motivated and easily interpreted representation of a SMBBH population, and did in fact perform better than both the $l_{\max} \leq 2$ and $L_{\max} \leq 2$ models.

Constraining deviations from isotropy in the nanohertz GWB is an enticing goal for the era beyond initial PTA detection. It is possible that as the GW signal significance grows in our data, loud binaries will first emerge from the background ensemble in excess angular power searches,

before eventually having their binary models constrained. As such, the fast and flexible techniques presented here form an important link in the chain of scientific milestones for PTAs, helping to kick-start single-source localization efforts even before other binary properties are known.

There are a variety of potential extensions to this work that build from assumptions we have made.

- (i) We have implicitly assumed a model for the GW angular power distribution that is the same at all frequencies. But this will not necessarily be the case in reality. In fact, the coalescence timescale of SMBBHs in the PTA band can be $\sim \text{Myrs}$, implying almost negligible evolution of the system within our observational window. For all intents and purposes, these are monochromatic signals. Thus one or several bright binaries may dominate over others in one frequency bin, but not in another, giving variable power maps across frequency. Indeed, as one moves toward higher frequencies in the PTA band, sources become sparser, implying a greater prospect for individual-source and anisotropic signatures. Adapting our framework for this case is trivial, simply requiring frequency-indexed overlap reduction functions.
- (ii) The SMBBH sky map was generated using the inclination-averaged strain of each binary, for which we then assumed that the power maps in $+$ and \times polarizations were equal. However, differences in these power polarization maps encode information, and any imbalance may indicate the presence of a bright binary dominating in a certain pixel. Relaxing our assumption of equal power in each polarization may further help in recognizing bright single sources, and can be easily achieved within our framework by allowing the polarization maps to be independently modeled.
- (iii) We have considered a statistically representative SMBBH population in order to produce a realistic GW sky map, however different populations will produce variations in anisotropy. Our population has several bright (somewhat clustered) sources that we were able to localize, but we did not consider a population where one binary overwhelmingly dominated the entire strain budget (which can sometimes happen). From a single-source detection standpoint, the population we considered may seem somewhat pessimistic, but this prospect varies significantly from realization to realization. In future work we will assess a range of population realizations to consider how well different models perform as a function of the power concentration in the loudest sources. It is likely that the single pixel modeling will be the most preferred model in a realization where a single source really sticks out.

In the spirit of open science, our code to rapidly compute PTA overlap reduction functions from a given angular power distribution is available at [105].

ACKNOWLEDGMENTS

We thank Sharan Banagiri, Vuk Mandic, Joe Romano, Eric Thrane, Ethan Payne, and Marc Kamionkowski, as well as colleagues in the NANOGrav collaboration and the International Pulsar Timing Array, for useful discussions. Much of this work was performed at the Jet Propulsion Laboratory, where S. R. T. was supported by appointment to the NASA Postdoctoral Program, administered by Oak Ridge Associated Universities and the Universities Space Research Association through a contract with NASA. S. R. T. was also supported by the NANOGrav NSF Physics Frontier Center Award No. 1430284. This work was supported in part by National Science Foundation Grant No. PHYS-1066293 and by the hospitality of the Aspen Center for Physics. A. S. is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program ERC-2018-COG under Grant Agreement No. 818691 (B Massive). A majority of the computational work was performed on the Nemo cluster at UWM supported by NSF Grant No. 0923409. Some of the results in this paper have been derived using the HEALPix [106] package.

APPENDIX A: DECOMPOSING POWER COEFFICIENTS c_{lm} IN TERMS OF SQUARE-ROOT POWER COEFFICIENTS a_{LM}

The calculation below requires Wigner-3j identities which are most easily expressed with complex spherical harmonics. The real spherical harmonics are merely a linear combination of these complex spherical harmonics, so the qualitative results are the same in both.

For the power decomposition in spherical harmonics we have

$$P(\hat{\Omega}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{lm} Y_l^m(\hat{\Omega}), \quad (\text{A1})$$

while for the square-root power decomposition in spherical harmonics we have

$$P(\hat{\Omega})^{1/2} = \sum_{L=0}^{\infty} \sum_{M=-L}^L a_{LM} Y_L^M(\hat{\Omega}), \quad (\text{A2})$$

where the notation Y_l^m, Y_L^M indicates the complex spherical harmonics.

Squaring Eq. (A2) and equating to Eq. (A1) gives

$$\begin{aligned} P(\hat{\Omega}) &= \sum_{LM} \sum_{L'M'} a_{LM} a_{L'M'} Y_L^M(\hat{\Omega}) Y_{L'}^{M'}(\hat{\Omega}) \\ &= \sum_{lm} c_{lm} Y_l^m(\hat{\Omega}). \end{aligned} \quad (\text{A3})$$

Thus, solving for c_{lm} gives

$$\begin{aligned} c_{lm} &= \sum_{LM} \sum_{L'M'} a_{LM} a_{L'M'} \int_{S^2} d\hat{\Omega} Y_l^{m*}(\hat{\Omega}) Y_L^M(\hat{\Omega}) Y_{L'}^{M'}(\hat{\Omega}) \\ &= (-1)^m \sum_{LM} \sum_{L'M'} a_{LM} a_{L'M'} \int_{S^2} d\hat{\Omega} Y_l^m(\hat{\Omega}) Y_L^M(\hat{\Omega}) Y_{L'}^{M'}(\hat{\Omega}), \end{aligned} \quad (\text{A4})$$

where the triple spherical-harmonic integral in Eq. (A4) can be expressed as a product of Wigner-3j symbols with the following identity:

$$\int_{S^2} d\hat{\Omega} Y_l^m(\hat{\Omega}) Y_L^M(\hat{\Omega}) Y_{L'}^{M'}(\hat{\Omega}) = \sqrt{\frac{(2l+1)(2L+1)(2L'+1)}{4\pi}} \begin{pmatrix} l & L & L' \\ -m & M & M' \end{pmatrix} \begin{pmatrix} l & L & L' \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{A5})$$

Hence, the power coefficients c_{lm} can be written as the following (infinite) summation over a_{LM} terms:

$$c_{lm} = (-1)^m \sum_{L=0}^{\infty} \sum_{M=-L}^L \sum_{L'=0}^{\infty} \sum_{M'=-L'}^{L'} a_{LM} a_{L'M'} \times \sqrt{\frac{(2l+1)(2L+1)(2L'+1)}{4\pi}} \begin{pmatrix} l & L & L' \\ -m & M & M' \end{pmatrix} \begin{pmatrix} l & L & L' \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{A6})$$

Equation (A6) shows that even a low order coefficient of power anisotropy may need contributions from higher orders in the square root of power, such that $L_{\max} > l_{\max}$. We investigate this with a numerical example, where we create an $N_{\text{side}} = 256$ HEALPix map of GWB power [shown in Fig. 7(a)] containing anisotropy up to and including $l = 2$, such that $P(\hat{\Omega}) = 1 + \cos \theta + 0.9 \times (3 \cos^2 \theta - 1)$. We also take the square root of this map, and use HEALPix functions to compute the angular power spectra of both $P(\hat{\Omega})$ and $P(\hat{\Omega})^{1/2}$. These spectra are shown in Fig. 7(b), where we see that $P(\hat{\Omega})$ contains anisotropic structure up to and including $l = 2$, but higher degrees drop to zero (within numerical precision). On the other hand, $P(\hat{\Omega})^{1/2}$ has anisotropic structure in degrees beyond quadrupole ($L > 2$) which are needed for a complete reconstruction of $P(\hat{\Omega})$. However, in practice we can achieve an adequate approximation of $P(\hat{\Omega})$ by squaring $P(\hat{\Omega})^{1/2}$ with $L \leq 2$. This is shown in Fig. 7(c), where we see the normalized match (inner product) between the original power map and a reconstruction

from the squaring of a restricted L decomposition of $P(\hat{\Omega})^{1/2}$. For $L_{\max} = 2$ the match is better than 98%.

APPENDIX B: PRODUCT SPACE SAMPLING

This technique recasts Bayesian model selection as a parameter estimation problem [18,107–109]. We define a hypermodel, \mathcal{H} , whose parameter space is the Cartesian product of the individual spaces of all submodels under consideration, along with an additional model indexing parameter, n . This model indexing parameter is discrete, but continuous sampling methods can simply be applied to a variable that is cast to an integer in the likelihood, or for which we define behavior based on parameter ranges.

The practical implementation of this is rather straightforward. At a given iteration in the sampling process we cast the model indexing parameter to an integer, which then indicates the “active” submodel for the likelihood evaluation. The hypermodel parameter space, θ_* , is partitioned

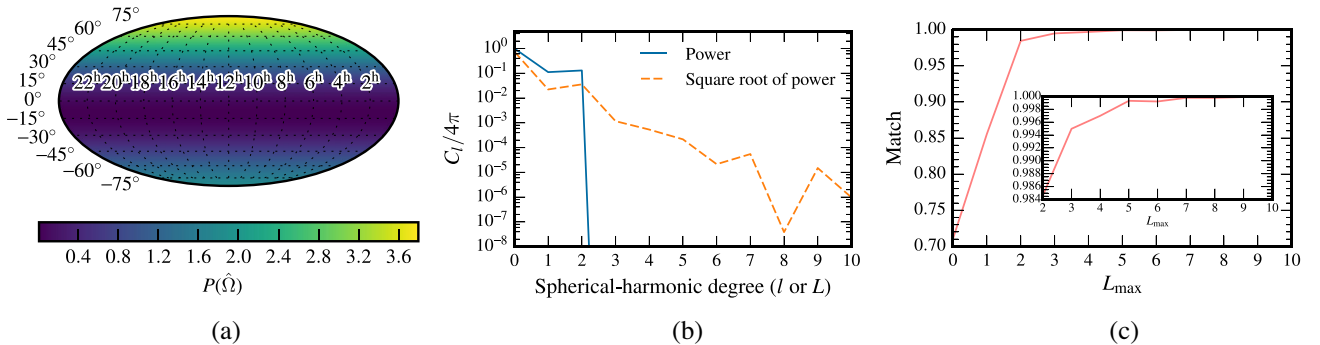


FIG. 7. (a) The tested power map with monopole, dipole, and quadrupole anisotropies. It is fully described by $P(\hat{\Omega}) = 1 + \cos \theta + 0.9 \times (3 \cos^2 \theta - 1)$. (b) The angular power spectra of $P(\hat{\Omega})$ and $P(\hat{\Omega})^{1/2}$. (c) The match between the original power map and a reconstruction from the squaring of a restricted L decomposition of $P(\hat{\Omega})^{1/2}$.

into the submodel spaces, and the corresponding parameters of the active submodel are passed to the relevant likelihood function. Thus the parameters of the inactive submodels do not contribute to, and are not constrained by, the active likelihood function. As sampling proceeds, the model indexing parameter will switch between all submodels to the end result that the relative fraction of sampling iterations spent in each submodel provides an estimate of the posterior odds ratio.

This result has been derived in other places (see e.g., [109]), but we reiterate it here for completeness. Consider constructing the marginalized posterior distribution of the model indexing parameter from the output of Markov chain Monte Carlo (MCMC) sampling:

$$\begin{aligned} p(n|\mathbf{d}, \mathcal{H}_*) &= \int p(\theta_*, n|\mathbf{d}, \mathcal{H}_*) d\theta \\ &= \frac{1}{\mathcal{Z}_*} \int p(\mathbf{d}|\theta_*, n, \mathcal{H}_*) p(\theta_*, n|\mathcal{H}_*) d\theta, \end{aligned} \quad (\text{B1})$$

where \mathcal{Z}_* is the hypermodel evidence. For a given n the hypermodel parameter space is partitioned into active, θ_n , and inactive, $\theta_{\bar{n}}$, parameters, $\theta_* = \{\theta_n, \theta_{\bar{n}}\}$, where the likelihood $p(\mathbf{d}|\theta_*, n, \mathcal{H}_*)$ is independent of the inactive parameters. The prior term can then be factorized:

$$p(\theta_*, n|\mathcal{H}_*) = p(\theta_n|\mathcal{H}_n) p(\theta_{\bar{n}}|\mathcal{H}_{\bar{n}}) p(n|\mathcal{H}_*) \quad (\text{B2})$$

such that

$$\begin{aligned} p(n|\mathbf{d}, \mathcal{H}_*) &= \frac{p(n|\mathcal{H}_*)}{\mathcal{Z}_*} \int p(\mathbf{d}|\theta_n, \mathcal{H}_n) p(\theta_n|\mathcal{H}_n) d\theta_n \\ &= \frac{p(n|\mathcal{H}_*)}{\mathcal{Z}_*} \mathcal{Z}_n, \end{aligned} \quad (\text{B3})$$

where \mathcal{Z}_n is the evidence for submodel n , and we have implicitly marginalized over inactive parameters since they only appear in the prior term $p(\theta_{\bar{n}}|\mathcal{H}_{\bar{n}})$ which simply integrates to unity. Thus the posterior odds ratios between two models is given by

$$\mathcal{P}_{12} = \frac{p(n_1|\mathcal{H}_*) \mathcal{Z}_1}{p(n_2|\mathcal{H}_*) \mathcal{Z}_2} = \frac{p(n_1|\mathbf{d}, \mathcal{H}_*)}{p(n_2|\mathbf{d}, \mathcal{H}_*)}, \quad (\text{B4})$$

where the hypermodel evidence cancels in this ratio of the two submodels. We add an additional refinement to this method, in that we perform a pilot run designed to provide an initial estimate of the posterior odds ratio, $\tilde{\mathcal{P}}_{12}$, then follow this up with a focused run for an improved estimate. In the focused run, we weight model 1 by $1/(1 + \tilde{\mathcal{P}}_{12})$ and model 2 by $\tilde{\mathcal{P}}_{12}/(1 + \tilde{\mathcal{P}}_{12})$. This has the benefit of improving mixing between models when the data significantly favor one model over another. The posterior odds ratio from the focused run is then reweighted to provide an improved estimate of \mathcal{P}_{12} . We estimate the uncertainty using methods for reversible jump MCMC [110], as the approaches are analogous. This product-space estimator of the posterior odds ratio is simple to implement, applicable to high dimensional parameter spaces, and allows direct model comparison.

-
- [1] J. H. Taylor and J. M. Weisberg, *Astrophys. J.* **253**, 908 (1982).
 - [2] J. Taylor, A. Wolszczan, T. Damour, and J. Weisberg, *Nature (London)* **355**, 132 (1992).
 - [3] M. Kramer, I. H. Stairs, R. N. Manchester, M. A. McLaughlin, A. G. Lyne, R. D. Ferdman, M. Burgay, D. R. Lorimer, A. Possenti, N. D'Amico, J. M. Sarkissian, G. B. Hobbs, J. E. Reynolds, P. C. C. Freire, and F. Camilo, *Science* **314**, 97 (2006).
 - [4] A. M. Archibald, N. V. Gusinskaia, J. W. T. Hessels, A. T. Deller, D. L. Kaplan, D. R. Lorimer, R. S. Lynch, S. M. Ransom, and I. H. Stairs, *Nature (London)* **559**, 73 (2018).
 - [5] R. S. Foster and D. C. Backer, *Astrophys. J.* **361**, 300 (1990).
 - [6] F. B. Estabrook and H. D. Wahlquist, *Gen. Relativ. Gravit.* **6**, 439 (1975).
 - [7] M. V. Sazhin, *Sov. Ast.* **22**, 36 (1978).
 - [8] S. Detweiler, *Astrophys. J.* **234**, 1100 (1979).
 - [9] M. Kramer and D. J. Champion, *Classical Quantum Gravity* **30**, 224009 (2013).
 - [10] R. van Haasteren, Y. Levin, G. H. Janssen, K. Lazaridis, M. Kramer, B. W. Stappers, G. Desvignes, M. B. Purver, A. G. Lyne, R. D. Ferdman, A. Jessner, I. Cognard, G. Theureau, N. D'Amico, A. Possenti, M. Burgay, A. Corongiu, J. W. T. Hessels, R. Smits, and J. P. W. Verbiest, *Mon. Not. R. Astron. Soc.* **414**, 3117 (2011).
 - [11] L. Lentati *et al.*, *Mon. Not. R. Astron. Soc.* **453**, 2577 (2015).
 - [12] S. Babak *et al.*, *Mon. Not. R. Astron. Soc.* **455**, 1665 (2016).
 - [13] M. A. McLaughlin, *Classical Quantum Gravity* **30**, 224008 (2013).
 - [14] P. B. Demorest *et al.*, *Astrophys. J.* **762**, 94 (2013).
 - [15] Z. Arzoumanian *et al.* (NANOGrav Collaboration), *Astrophys. J.* **794**, 141 (2014).
 - [16] Z. Arzoumanian *et al.* (NANOGrav Collaboration), *Astrophys. J.* **810**, 150 (2015).
 - [17] Z. Arzoumanian *et al.* (NANOGrav Collaboration), *Astrophys. J.* **821**, 13 (2016).
 - [18] Z. Arzoumanian *et al.* (NANOGrav Collaboration), *Astrophys. J.* **859**, 47 (2018).
 - [19] K. Aggarwal *et al.* (NANOGrav Collaboration), *Astrophys. J.* **880**, 116 (2019).

- [20] K. Aggarwal *et al.*, *Astrophys. J.* **889**, 38 (2020).
- [21] R. N. Manchester *et al.*, *Pub. Astron. Soc. Aust.* **30**, e017 (2013).
- [22] G. Hobbs, *Classical Quantum Gravity* **30**, 224007 (2013).
- [23] R. M. Shannon, V. Ravi, W. A. Coles, G. Hobbs, M. J. Keith, R. N. Manchester, J. S. B. Wyithe, M. Bailes, N. D. R. Bhat, S. Burke-Spolaor, J. Khoo, Y. Levin, S. Osłowski, J. M. Sarkissian, W. van Straten, J. P. W. Verbiest, and J. B. Wang, *Science* **342**, 334 (2013).
- [24] X. J. Zhu, G. Hobbs, L. Wen, W. A. Coles, J. B. Wang, R. M. Shannon, R. N. Manchester, M. Bailes, N. D. R. Bhat, S. Burke-Spolaor, S. Dai, M. J. Keith, M. Kerr, Y. Levin, D. R. Madison, S. Osłowski, V. Ravi, L. Toomey, and W. van Straten, *Mon. Not. R. Astron. Soc.* **444**, 3709 (2014).
- [25] J. B. Wang, G. Hobbs, W. Coles, R. M. Shannon, X. J. Zhu, D. R. Madison, M. Kerr, V. Ravi, M. J. Keith, R. N. Manchester, Y. Levin, M. Bailes, N. D. R. Bhat, S. Burke-Spolaor, S. Dai, S. Osłowski, W. van Straten, L. Toomey, N. Wang, and L. Wen, *Mon. Not. R. Astron. Soc.* **446**, 1657 (2015).
- [26] R. M. Shannon *et al.*, *Science* **349**, 1522 (2015).
- [27] D. R. Madison *et al.*, *Mon. Not. R. Astron. Soc.* **455**, 3662 (2016).
- [28] J. P. W. Verbiest *et al.*, *Mon. Not. R. Astron. Soc.* **458**, 1267 (2016).
- [29] B. B. P. Perera *et al.*, *Mon. Not. R. Astron. Soc.* **490**, 4666 (2019).
- [30] B. C. Joshi, P. Arumugasamy, M. Bagchi, D. Bandyopadhyay, A. Basu, N. Dhand a Batra, S. Bethapudi, A. Choudhary, K. De, L. Dey, A. Gopakumar, Y. Gupta, M. A. Krishnakumar, Y. Maan, P. K. Manoharan, A. Naidu, R. Nandi, D. Pathak, M. Surnis, and A. Susobhanan, *J. Astrophys. Astron.* **39**, 51 (2018).
- [31] K. J. Lee, Prospects of gravitational wave detection using pulsar timing array for chinese future telescopes, in *Frontiers in Radio Astronomy and FAST Early Sciences Symposium 2015*, Astronomical Society of the Pacific Conference Series Vol. 502, edited by L. Qain and D. Li (Astronomical Society of the Pacific, 2016), p. 19, <http://aspbooks.org/custom/publications/paper/502-0019.html>.
- [32] M. Bailes *et al.*, *Proc. Sci.*, MeerKAT2016 (2018) 011 [arXiv:1803.07424].
- [33] C. Ng, *IAU Symp.* **13**, 179 (2017).
- [34] M. C. Begelman, R. D. Blandford, and M. J. Rees, *Nature (London)* **287**, 307 (1980).
- [35] J. S. B. Wyithe and A. Loeb, *Astrophys. J.* **590**, 691 (2003).
- [36] A. H. Jaffe and D. C. Backer, *Astrophys. J.* **583**, 616 (2003).
- [37] A. Sesana, A. Vecchio, and C. N. Colacino, *Mon. Not. R. Astron. Soc.* **390**, 192 (2008).
- [38] J. Kormendy and D. Richstone, *Annu. Rev. Astron. Astrophys.* **33**, 581 (1995).
- [39] J. Magorrian, S. Tremaine, D. Richstone, R. Bender, G. Bower, A. Dressler, S. M. Faber, K. Gebhardt, R. Green, C. Grillmair, J. Kormendy, and T. Lauer, *Astron. J.* **115**, 2285 (1998).
- [40] L. Ferrarese and D. Merritt, *Astrophys. J.* **539**, L9 (2000).
- [41] K. Gebhardt, R. Bender, G. Bower, A. Dressler, S. M. Faber, A. V. Filippenko, R. Green, C. Grillmair, L. C. Ho, J. Kormendy, T. R. Lauer, J. Magorrian, J. Pinkney, D. Richstone, and S. Tremaine, *Astrophys. J.* **539**, L13 (2000).
- [42] S. Burke-Spolaor, S. R. Taylor, M. Charisi, T. Dolch, J. S. Hazboun, A. M. Holgado, L. Z. Kelley, T. J. W. Lazio, D. R. Madison, N. McMann, C. M. F. Mingarelli, A. Rasskazov, X. Siemens, J. J. Simon, and T. L. Smith, *Astron. Astrophys. Rev.* **27**, 5 (2019).
- [43] S. Babak and A. Sesana, *Phys. Rev. D* **85**, 044034 (2012).
- [44] A. Petiteau, S. Babak, A. Sesana, and M. de Araújo, *Phys. Rev. D* **87**, 064036 (2013).
- [45] L. Boyle and U.-L. Pen, *Phys. Rev. D* **86**, 124028 (2012).
- [46] P. A. Rosado, A. Sesana, and J. Gair, *Mon. Not. R. Astron. Soc.* **451**, 2417 (2015).
- [47] L. Z. Kelley, L. Blecha, L. Hernquist, A. Sesana, and S. R. Taylor, *Mon. Not. R. Astron. Soc.* **477**, 964 (2018).
- [48] A. Sesana, *Mon. Not. R. Astron. Soc.* **433**, L1 (2013).
- [49] A. Sesana, A. Vecchio, and M. Volonteri, *Mon. Not. R. Astron. Soc.* **394**, 2255 (2009).
- [50] A. Sesana and A. Vecchio, *Phys. Rev. D* **81**, 104008 (2010).
- [51] K. J. Lee, N. Wex, M. Kramer, B. W. Stappers, C. G. Bassa, G. H. Janssen, R. Karuppusamy, and R. Smits, *Mon. Not. R. Astron. Soc.* **414**, 3251 (2011).
- [52] J. A. Ellis, X. Siemens, and J. D. E. Creighton, *Astrophys. J.* **756**, 175 (2012).
- [53] J. A. Ellis, *Classical Quantum Gravity* **30**, 224004 (2013).
- [54] S. Taylor, J. Ellis, and J. Gair, *Phys. Rev. D* **90**, 104028 (2014).
- [55] S. R. Taylor, E. A. Huerta, J. R. Gair, and S. T. McWilliams, *Astrophys. J.* **817**, 70 (2016).
- [56] C. M. F. Mingarelli, T. Sidery, I. Mandel, and A. Vecchio, *Phys. Rev. D* **88**, 062005 (2013).
- [57] S. R. Taylor and J. R. Gair, *Phys. Rev. D* **88**, 084001 (2013).
- [58] C. M. F. Mingarelli, T. J. W. Lazio, A. Sesana, J. E. Greene, J. A. Ellis, C.-P. Ma, S. Croft, S. Burke-Spolaor, and S. R. Taylor, *Nat. Astron.* **1**, 886 (2017).
- [59] E. Roebber and G. Holder, *Astrophys. J.* **835**, 21 (2017).
- [60] R. W. Hellings and G. S. Downs, *Astrophys. J. Lett.* **265**, L39 (1983).
- [61] E. Thrane, S. Ballmer, J. D. Romano, S. Mitra, D. Talukder, S. Bose, and V. Mandic, *Phys. Rev. D* **80**, 122002 (2009).
- [62] J. Gair, J. D. Romano, S. Taylor, and C. M. F. Mingarelli, *Phys. Rev. D* **90**, 082001 (2014).
- [63] S. R. Taylor *et al.* (EPTA Collaboration), *Phys. Rev. Lett.* **115**, 041101 (2015).
- [64] C. Conneely, A. H. Jaffe, and C. M. F. Mingarelli, *Mon. Not. R. Astron. Soc.* **487**, 562 (2019).
- [65] S. C. Hotinli, M. Kamionkowski, and A. H. Jaffe, *Open J. Astrophys.* **2**, 8 (2019).
- [66] N. J. Cornish and R. van Haasteren, [arXiv:1406.4511](https://arxiv.org/abs/1406.4511).
- [67] Y. Ali-Haïmoud, T. L. Smith, and C. M. F. Mingarelli, [arXiv:2006.14570](https://arxiv.org/abs/2006.14570).
- [68] B. Bécsy and N. J. Cornish, *Classical Quantum Gravity* **37**, 135011 (2020).
- [69] J. M. Goldstein, A. Sesana, A. M. Holgado, and J. Veitch, *Mon. Not. R. Astron. Soc.* **485**, 248 (2019).
- [70] M. J. Graham, S. G. Djorgovski, D. Stern, A. J. Drake, A. A. Mahabal, C. Donalek, E. Glikman, S. Larson, and E. Christensen, *Mon. Not. R. Astron. Soc.* **453**, 1562 (2015).

- [71] F. A. Jenet, A. Lommen, S. L. Larson, and L. Wen, *Astrophys. J.* **606**, 799 (2004).
- [72] D. J. D’Orazio, Z. Haiman, P. Duffell, B. D. Farris, and A. I. MacFadyen, *Mon. Not. R. Astron. Soc.* **452**, 2540 (2015).
- [73] M. Charisi, I. Bartos, Z. Haiman, A. M. Price-Whelan, M. J. Graham, E. C. Bellm, R. R. Laher, and S. Márka, *Mon. Not. R. Astron. Soc.* **463**, 2145 (2016).
- [74] T. Liu, S. Gezari, M. Ayers, W. Burgett, K. Chambers, K. Hodapp, M. E. Huber, R. P. Kudritzki, N. Metcalfe, J. Tonry, R. Wainscoat, and C. Waters, *Astrophys. J.* **884**, 36 (2019).
- [75] A. M. Holgado, A. Sesana, A. Sandrinelli, S. Covino, A. Treves, X. Liu, and P. Ricker, *Mon. Not. R. Astron. Soc.* **481**, L74 (2018).
- [76] A. Sesana, Z. Haiman, B. Kocsis, and L. Z. Kelley, *Astrophys. J.* **856**, 42 (2018).
- [77] X.-J. Zhu and E. Thrane, [arXiv:2004.10944](https://arxiv.org/abs/2004.10944).
- [78] J. Simon, A. Polin, A. Lommen, B. Stappers, L. S. Finn, F. A. Jenet, and B. Christy, *Astrophys. J.* **784**, 60 (2014).
- [79] M. Anholm, S. Ballmer, J. D. E. Creighton, L. R. Price, and X. Siemens, *Phys. Rev. D* **79**, 084030 (2009).
- [80] L. G. Book and É. É. Flanagan, *Phys. Rev. D* **83**, 024024 (2011).
- [81] E. E. Flanagan, *Phys. Rev. D* **48**, 2389 (1993).
- [82] B. Allen and A. C. Ottewill, *Phys. Rev. D* **56**, 545 (1997).
- [83] E. Payne, S. Banagiri, P. Lasky, and E. Thrane, [arXiv:2006.11957](https://arxiv.org/abs/2006.11957).
- [84] S. W. Ballmer, *Classical Quantum Gravity* **23**, S179 (2006).
- [85] S. Mitra, S. Dhurandhar, T. Souradeep, A. Lazzarini, V. Mandic, S. Bose, and S. Ballmer, *Phys. Rev. D* **77**, 042002 (2008).
- [86] N. J. Cornish and A. Sesana, *Classical Quantum Gravity* **30**, 224005 (2013).
- [87] Z. Arzoumanian *et al.* (The NANOGrav Collaboration), *Astrophys. J.* **813**, 65 (2015).
- [88] M. Rajagopal and R. W. Romani, *Astrophys. J.* **446**, 543 (1995).
- [89] E. S. Phinney, [arXiv:astro-ph/0108028](https://arxiv.org/abs/astro-ph/0108028).
- [90] L. Lentati, P. Alexander, M. P. Hobson, S. Taylor, J. Gair, S. T. Balan, and R. van Haasteren, *Phys. Rev. D* **87**, 104021 (2013).
- [91] R. van Haasteren and M. Vallisneri, *Phys. Rev. D* **90**, 104012 (2014).
- [92] R. van Haasteren and M. Vallisneri, *Mon. Not. R. Astron. Soc.* **446**, 1170 (2015).
- [93] W. Coles, G. Hobbs, D. J. Champion, R. N. Manchester, and J. P. W. Verbiest, *Mon. Not. R. Astron. Soc.* **418**, 561 (2011).
- [94] A. Muzzin, D. Marchesini, M. Stefanon, M. Franx, H. J. McCracken, B. Milvang-Jensen, J. S. Dunlop, J. P. U. Fynbo, G. Brammer, I. Labbé, and P. G. van Dokkum, *Astrophys. J.* **777**, 18 (2013).
- [95] K. Bundy, M. Fukugita, R. S. Ellis, T. A. Targett, S. Belli, and T. Kodama, *Astrophys. J.* **697**, 1369 (2009).
- [96] J. M. Lotz, P. Jonsson, T. J. Cox, and J. R. Primack, *Mon. Not. R. Astron. Soc.* **404**, 575 (2010).
- [97] N. Häring and H.-W. Rix, *Astrophys. J.* **604**, L89 (2004).
- [98] M. Faisal Alam *et al.*, [arXiv:2005.06490](https://arxiv.org/abs/2005.06490).
- [99] M. Faisal Alam *et al.*, [arXiv:2005.06495](https://arxiv.org/abs/2005.06495).
- [100] S. Ransom, A. Brazier, S. Chatterjee, T. Cohen, J. M. Cordes, M. E. DeCesar, P. B. Demorest, J. S. Hazboun, M. T. Lam, R. S. Lynch, M. A. McLaughlin, S. M. Ransom, X. Siemens, S. R. Taylor, and S. J. Vigeland, in *BAAS* (AAS Publishing, 2019), Vol. 51, p. 195.
- [101] S. R. Taylor, M. Vallisneri, J. A. Ellis, C. M. F. Mingarelli, T. J. W. Lazio, and R. van Haasteren, *Astrophys. J.* **819**, L6 (2016).
- [102] N. J. Cornish and L. Sampson, *Phys. Rev. D* **93**, 104047 (2016).
- [103] S. R. Taylor, L. Lentati, S. Babak, P. Brem, J. R. Gair, A. Sesana, and A. Vecchio, *Phys. Rev. D* **95**, 042002 (2017).
- [104] X. Siemens, J. Ellis, F. Jenet, and J. D. Romano, *Classical Quantum Gravity* **30**, 224015 (2013).
- [105] <https://github.com/stevertaylor/gworf>.
- [106] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, *Astrophys. J.* **622**, 759 (2005).
- [107] B. P. Carlin and S. Chib, *J. R. Stat. Soc. Ser. B* **57**, 473 (1995).
- [108] S. J. Godsill, *J. Comput. Graph. Stat.* **10**, 230 (2001).
- [109] S. Hee, W. J. Handley, M. P. Hobson, and A. N. Lasenby, *Mon. Not. R. Astron. Soc.* **455**, 2461 (2016).
- [110] N. J. Cornish and T. B. Littenberg, *Classical Quantum Gravity* **32**, 135012 (2015).